SEQUENCE LISTING

<110> Johnson, Jeffrey D.
Palma, John F.
Schweitzer, Anthony C.
Blume, John E.
Metabolex, Inc.

<120> A Pancreatic Islet Transcription Factor and Uses Thereof

<130> 016325-013510US

<140> US 10/533,593
<141> 2005-05-02

<150> US 60/425,968
<151> 2002-11-13

<150> WO PCT/US03/36131
<151> 2003-11-13

<160> 42

<170> PatentIn Ver. 2.1

<210> 1
<211> 3048
<212> DNA
<213> Homo sapiens

<220>
<223> human IC-RFX islet transcription factor cDNA

<400> 1
tttctgcgct gagccagggc accccggagc ctgcggcctc cttccccgcc cctgcggccc 60
cgggtcccag ccccgccccg ccccgccccg ggctggggct ccgctgggga accggccgag 120
cggcgcgcgc ggaggtgtcc ggcggccagg aggatggcca aggtcccgga gctggaagac 180
accttcctgc aggcgcagcc tgcgccccaa ctgtccccgg ggatccagga agactgctgt 240
gtgcagctcc tgggcaaggg cttgctagtc tatccggaag aaacagtgta cctggcggcc 300
gaagggcagc ccgggggcga gcagggcggc ggggagaaag gcgaagaccc ggagctgccg 360
ggggcagtga aatcagaaat gcacttaaac aatggtaact tttcctctga agaagaggac 420
gccgacaacc acgacagcaa aaccaaagca gcggatcaat acctgtctca gaagaaaacc 480
atcacgcaga ttgtgaagga taaaaagaag cagacacagc tcacgctgca gtggcttgaa 540
gagaattaca ttgtatgtga aggagtttgc ttaccacggt gcattcttta tgcacactac 600
ttagatttct gtaggaaaga gaaattagag ccagcctgtg cggccacctt tggaaagaca 660
attcgccaga agtttcccct cctaacaaca aggcggcttg gaacaagagg ccattcaaag 720
tatcattact atgggattgg catcaaagag agcagtgcat attaccactc cgtttattct 780
ggaaagggct tgacaaggtt ttctggaagc aagctaaaga atgagggtgg cttcactcgt 840
aaatattcgc ttagctcaaa aactggaaca cttcttccag aattccccag cgctcaacac 900
cttgtatacc aaggatgcat ttctaaggac aaggttgata cgctcataat gatgtacaaa 960
actcactgcc agtgtatcct ggacaatgca attaatggaa actttgaaga gatccagcat 1020
tttttattac acttttggca aggaatgcct gaccatctcc ttcccctgct cgaaaatcct 1080
gttatcattg atattttctg tgtttgtgac tcaattcttt ataaggttct tacagatgta 1140
ctcattcctg caacaatgca agaaatgcct gaaagcttat tagcagacat aagaaatttt 1200
gctaaaaatt gggaacagtg ggttgtttca tccttggaaa acttgccaga agctctaact 1260
gacaagaaaa tacctattgt gcgaagattt gtatcttctc tgaaacgaca aacatctttc 1320
ttacatcttg cccagattgc cagaccagct ctctttgacc agcatgtcgt taattctatg 1380
gtgtctgata ttgaaagggt tgatttgaac agcattggct ctcaagccct tcttaccatt 1440
tcaggcagca cagacactga atctggtatc tacactgaac atgactctat cactgtgttc 1500
caagaactga aggatctcct taagaagaat gccactgtgg aggcttttat tgaatggttg 1560

```
gatactgtgg tagaacagag agttattaag accagcaaac aaaatggaag gtcattaaag 1620
aagagagctc aagactttct gttaaagtgg agtttttttg gtgctcgagt aatgcataat 1680
ctcaccttga acaatgcatc cagttttggt tcttttcatt tgattcgaat gcttctcgat 1740
gaatacattc tcctggccat ggagacccag tttaataatg acaaagagca ggagttacag 1800
aatttattgg acaagtatat gaagaattca gatgcgagta aagctgcttt cactgcttct 1860
ccgagttcat gctttctggc caaccgtaat aaagggagca tggtttccag cgacgctgtg 1920
aagaatgaaa gccacgtgga gacaacctat ctccctctgc catccagtca acctggaggc 1980
ctaggccctg ctctgcacca gttccctgct gggaacacag acaacatgcc gctcacaggt 2040
caaatggagc tttcacagat tgctggtcat ctgatgacac cacccatttc tccagccatg 2100
gcaagccgag gaagtgtcat taaccaagga ccaatggcag ggaggccccc aagtgtgggc 2160
ccagtactgt cagctccatc acactgctcc acatacccag agcccattta tcccactctc 2220
cctcaagcca atcatgactt ttatagcacc agctctaact accagactgt gtttagggca 2280
cagccccact ccacatcagg actctatcct catcacaccg agcatggtcg atgcatggct 2340
tggactgaac agcagctttc aagagacttc ttcagtggca gctgtgcggg gtctccatat 2400
aactcccggc caccgtctag ctatggccca tccctgcaag cccaggattc acacaatatg 2460
cagttttaa atacaggaag cttcaatttc ttgagcaaca caggagctgc cagctgccaa 2520
ggagcaacac tgcctcctaa ttcaccaaat ggatactatg gaagcaacat aaactaccca 2580
gagtctcaca ggctcggatc aatggtgaat cagcacgttt ctgtcatcag cagcattcgt 2640
tcactgcccc cctacagtga catccacgat ccacttaaca ttttagatga cagtggtaga 2700
aaacagacca gctcgtttta cacagacaca tcatctccag ttgcatgtcg aactccagtc 2760
ctagcttcca gtttgcaaac cccaattcct tcttcctcat cccaatgtat gtatggaact 2820
tccaaccagt atccagctca agaaaccctg gactcccatg gaacaagcag tagagaaatg 2880
gtgtcctctt taccacctat caacactgtg ttcatgggaa cagcagctgg aggcacttaa 2940
accaccaatg tgggaggggg tgctaaaact ttaaaaaaaa tctctactgt gcaaatatca 3000
ttattcactc agacttccat aagagtaaat aaaaaatgaa tatgcagt               3048
```

<210> 2
<211> 928
<212> PRT
<213> Homo sapiens

<220>
<223> human IC-RFX islet transcription factor

<400> 2

```
Met Ala Lys Val Pro Glu Leu Glu Asp Thr Phe Leu Gln Ala Gln Pro
  1               5                  10                  15

Ala Pro Gln Leu Ser Pro Gly Ile Gln Glu Asp Cys Cys Val Gln Leu
             20                  25                  30

Leu Gly Lys Gly Leu Leu Val Tyr Pro Glu Glu Thr Val Tyr Leu Ala
         35                  40                  45

Ala Glu Gly Gln Pro Gly Gly Glu Gln Gly Gly Gly Glu Lys Gly Glu
     50                  55                  60

Asp Pro Glu Leu Pro Gly Ala Val Lys Ser Glu Met His Leu Asn Asn
 65                  70                  75                  80

Gly Asn Phe Ser Ser Glu Glu Glu Asp Ala Asp Asn His Asp Ser Lys
                 85                  90                  95

Thr Lys Ala Ala Asp Gln Tyr Leu Ser Gln Lys Lys Thr Ile Thr Gln
            100                 105                 110

Ile Val Lys Asp Lys Lys Lys Gln Thr Gln Leu Thr Leu Gln Trp Leu
        115                 120                 125
```

```
Glu Glu Asn Tyr Ile Val Cys Glu Gly Val Cys Leu Pro Arg Cys Ile
    130                 135                 140

Leu Tyr Ala His Tyr Leu Asp Phe Cys Arg Lys Glu Lys Leu Glu Pro
145                 150                 155                 160

Ala Cys Ala Ala Thr Phe Gly Lys Thr Ile Arg Gln Lys Phe Pro Leu
                165                 170                 175

Leu Thr Thr Arg Arg Leu Gly Thr Arg Gly His Ser Lys Tyr His Tyr
            180                 185                 190

Tyr Gly Ile Gly Ile Lys Glu Ser Ser Ala Tyr Tyr His Ser Val Tyr
        195                 200                 205

Ser Gly Lys Gly Leu Thr Arg Phe Ser Gly Ser Lys Leu Lys Asn Glu
    210                 215                 220

Gly Gly Phe Thr Arg Lys Tyr Ser Leu Ser Ser Lys Thr Gly Thr Leu
225                 230                 235                 240

Leu Pro Glu Phe Pro Ser Ala Gln His Leu Val Tyr Gln Gly Cys Ile
                245                 250                 255

Ser Lys Asp Lys Val Asp Thr Leu Ile Met Met Tyr Lys Thr His Cys
            260                 265                 270

Gln Cys Ile Leu Asp Asn Ala Ile Asn Gly Asn Phe Glu Glu Ile Gln
        275                 280                 285

His Phe Leu Leu His Phe Trp Gln Gly Met Pro Asp His Leu Leu Pro
    290                 295                 300

Leu Leu Glu Asn Pro Val Ile Ile Asp Ile Phe Cys Val Cys Asp Ser
305                 310                 315                 320

Ile Leu Tyr Lys Val Leu Thr Asp Val Leu Ile Pro Ala Thr Met Gln
                325                 330                 335

Glu Met Pro Glu Ser Leu Leu Ala Asp Ile Arg Asn Phe Ala Lys Asn
            340                 345                 350

Trp Glu Gln Trp Val Val Ser Ser Leu Glu Asn Leu Pro Glu Ala Leu
        355                 360                 365

Thr Asp Lys Lys Ile Pro Ile Val Arg Arg Phe Val Ser Ser Leu Lys
    370                 375                 380

Arg Gln Thr Ser Phe Leu His Leu Ala Gln Ile Ala Arg Pro Ala Leu
385                 390                 395                 400

Phe Asp Gln His Val Val Asn Ser Met Val Ser Asp Ile Glu Arg Val
                405                 410                 415

Asp Leu Asn Ser Ile Gly Ser Gln Ala Leu Leu Thr Ile Ser Gly Ser
            420                 425                 430

Thr Asp Thr Glu Ser Gly Ile Tyr Thr Glu His Asp Ser Ile Thr Val
        435                 440                 445
```

```
Phe Gln Glu Leu Lys Asp Leu Leu Lys Lys Asn Ala Thr Val Glu Ala
    450                 455                 460

Phe Ile Glu Trp Leu Asp Thr Val Val Glu Gln Arg Val Ile Lys Thr
465                 470                 475                 480

Ser Lys Gln Asn Gly Arg Ser Leu Lys Lys Arg Ala Gln Asp Phe Leu
                485                 490                 495

Leu Lys Trp Ser Phe Phe Gly Ala Arg Val Met His Asn Leu Thr Leu
            500                 505                 510

Asn Asn Ala Ser Ser Phe Gly Ser Phe His Leu Ile Arg Met Leu Leu
        515                 520                 525

Asp Glu Tyr Ile Leu Leu Ala Met Glu Thr Gln Phe Asn Asn Asp Lys
    530                 535                 540

Glu Gln Glu Leu Gln Asn Leu Leu Asp Lys Tyr Met Lys Asn Ser Asp
545                 550                 555                 560

Ala Ser Lys Ala Ala Phe Thr Ala Ser Pro Ser Ser Cys Phe Leu Ala
                565                 570                 575

Asn Arg Asn Lys Gly Ser Met Val Ser Ser Asp Ala Val Lys Asn Glu
            580                 585                 590

Ser His Val Glu Thr Thr Tyr Leu Pro Leu Pro Ser Ser Gln Pro Gly
        595                 600                 605

Gly Leu Gly Pro Ala Leu His Gln Phe Pro Ala Gly Asn Thr Asp Asn
    610                 615                 620

Met Pro Leu Thr Gly Gln Met Glu Leu Ser Gln Ile Ala Gly His Leu
625                 630                 635                 640

Met Thr Pro Pro Ile Ser Pro Ala Met Ala Ser Arg Gly Ser Val Ile
                645                 650                 655

Asn Gln Gly Pro Met Ala Gly Arg Pro Pro Ser Val Gly Pro Val Leu
            660                 665                 670

Ser Ala Pro Ser His Cys Ser Thr Tyr Pro Glu Pro Ile Tyr Pro Thr
        675                 680                 685

Leu Pro Gln Ala Asn His Asp Phe Tyr Ser Thr Ser Ser Asn Tyr Gln
    690                 695                 700

Thr Val Phe Arg Ala Gln Pro His Ser Thr Ser Gly Leu Tyr Pro His
705                 710                 715                 720

His Thr Glu His Gly Arg Cys Met Ala Trp Thr Glu Gln Gln Leu Ser
                725                 730                 735

Arg Asp Phe Phe Ser Gly Ser Cys Ala Gly Ser Pro Tyr Asn Ser Arg
            740                 745                 750

Pro Pro Ser Ser Tyr Gly Pro Ser Leu Gln Ala Gln Asp Ser His Asn
        755                 760                 765
```

```
Met Gln Phe Leu Asn Thr Gly Ser Phe Asn Phe Leu Ser Asn Thr Gly
    770                 775                 780

Ala Ala Ser Cys Gln Gly Ala Thr Leu Pro Pro Asn Ser Pro Asn Gly
785                 790                 795                 800

Tyr Tyr Gly Ser Asn Ile Asn Tyr Pro Glu Ser His Arg Leu Gly Ser
                805                 810                 815

Met Val Asn Gln His Val Ser Val Ile Ser Ser Ile Arg Ser Leu Pro
            820                 825                 830

Pro Tyr Ser Asp Ile His Asp Pro Leu Asn Ile Leu Asp Asp Ser Gly
        835                 840                 845

Arg Lys Gln Thr Ser Ser Phe Tyr Thr Asp Thr Ser Ser Pro Val Ala
    850                 855                 860

Cys Arg Thr Pro Val Leu Ala Ser Ser Leu Gln Thr Pro Ile Pro Ser
865                 870                 875                 880

Ser Ser Ser Gln Cys Met Tyr Gly Thr Ser Asn Gln Tyr Pro Ala Gln
                885                 890                 895

Glu Thr Leu Asp Ser His Gly Thr Ser Ser Arg Glu Met Val Ser Ser
            900                 905                 910

Leu Pro Pro Ile Asn Thr Val Phe Met Gly Thr Ala Ala Gly Gly Thr
        915                 920                 925
```

<210> 3
<211> 928
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:exemplary
IC-RFX consensus sequence

<220>
<221> MOD_RES
<222> (1)..(928)
<223> Xaa = any amino acid

<400> 3

```
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
  1               5                  10                  15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            20                  25                  30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        35                  40                  45

Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Gly Gly Xaa Xaa Xaa Xaa Xaa
    50                  55                  60

Xaa Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Glu Xaa Xaa Leu Asn Xaa
65                  70                  75                  80
```

Gly Xaa Xaa Xaa Ser Glu Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Ser Xaa
85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Xaa Xaa Lys Xaa Xaa Xaa Thr Xaa Xaa Thr Leu Gln Trp Leu
115 120 125

Glu Glu Asn Tyr Xaa Xaa Xaa Glu Gly Val Cys Leu Pro Arg Cys Xaa
130 135 140

Leu Tyr Xaa His Tyr Leu Asp Phe Cys Xaa Lys Xaa Xaa Xaa Xaa Pro
145 150 155 160

Xaa Xaa Ala Ala Xaa Phe Gly Lys Xaa Ile Arg Gln Xaa Phe Pro Xaa
165 170 175

Leu Thr Thr Arg Arg Leu Gly Thr Arg Gly Xaa Ser Lys Tyr His Tyr
180 185 190

Tyr Gly Ile Xaa Xaa Lys Glu Ser Ser Xaa Tyr Tyr Xaa Xaa Xaa Tyr
195 200 205

Ser Xaa Lys Gly Xaa Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa Lys Xaa Xaa
210 215 220

Xaa Xaa Xaa Thr Xaa Xaa Tyr Ser Xaa Xaa Ser Lys Xaa Gly Thr Leu
225 230 235 240

Leu Pro Glu Phe Pro Xaa Xaa Gln His Xaa Xaa Xaa Xaa Xaa Xaa Ile
245 250 255

Xaa Xaa Xaa Lys Val Xaa Thr Leu Ile Met Met Tyr Xaa Thr His Cys
260 265 270

Gln Xaa Ile Leu Asp Xaa Xaa Ile Xaa Xaa Asn Phe Xaa Glu Xaa Gln
275 280 285

Xaa Phe Leu Leu His Phe Trp Gln Gly Met Pro Asp His Xaa Leu Pro
290 295 300

Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp Ile Xaa Xaa Val Cys Asp Ser
305 310 315 320

Ile Leu Tyr Lys Xaa Xaa Xaa Asp Val Leu Ile Pro Xaa Xaa Xaa Gln
325 330 335

Glu Xaa Pro Xaa Ser Leu Xaa Xaa Xaa Ile Arg Xaa Phe Ala Lys Asn
340 345 350

Xaa Xaa Xaa Trp Xaa Xaa Xaa Ser Leu Xaa Asn Leu Pro Glu Xaa Leu
355 360 365

Xaa Xaa Lys Lys Ile Xaa Xaa Xaa Arg Arg Phe Xaa Xaa Xaa Leu Lys
370 375 380

Arg Gln Thr Ser Xaa Xaa His Leu Ala Gln Xaa Xaa Arg Xaa Xaa Leu
385 390 395 400

Xaa Xaa Xaa Xaa Xaa Val Xaa Xaa Met Xaa Xaa Asp Xaa Xaa Arg Val
405 410 415

Asp Leu Asn Ser Ile Xaa Xaa Gln Ala Leu Xaa Thr Xaa Xaa Xaa Ser
420 425 430

Xaa Asp Xaa Xaa Xaa Xaa Xaa Xaa Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa
435 440 445

Xaa Gln Glu Xaa Lys Xaa Leu Leu Xaa Xaa Xaa Xaa Xaa Xaa Glu Xaa
450 455 460

Xaa Ile Glu Trp Leu Asp Thr Xaa Val Xaa Gln Xaa Val Xaa Lys Xaa
465 470 475 480

Xaa Xaa Xaa Xaa Xaa Xaa Ser Leu Lys Lys Xaa Ala Gln Xaa Phe Leu
485 490 495

Leu Xaa Trp Ser Phe Phe Gly Xaa Arg Val Xaa Xaa Xaa Leu Thr Leu
500 505 510

Xaa Xaa Ala Xaa Ser Phe Gly Ser Phe His Leu Ile Arg Xaa Leu Xaa
515 520 525

Asp Xaa Tyr Xaa Leu Xaa Xaa Xaa Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
530 535 540

Xaa Gln Glu Xaa Xaa Asn Xaa Leu Xaa Xaa Xaa Met Lys Xaa Xaa Xaa
545 550 555 560

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
565 570 575

Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Ser Xaa Xaa Xaa Val Xaa Xaa Glu
580 585 590

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Leu Pro Ser Xaa Xaa Xaa Xaa
595 600 605

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
610 615 620

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
625 630 635 640

Xaa Xaa Pro Xaa Xaa Ser Pro Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
645 650 655

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Pro Xaa Xaa Xaa Xaa Xaa Xaa
660 665 670

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
675 680 685

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa Xaa Xaa
690 695 700

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
705 710 715 720

```
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                725                 730                 735

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Pro Xaa Xaa Xaa Xaa
            740                 745                 750

Xaa Pro Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        755                 760                 765

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Ser Asn Thr Gly
    770                 775                 780

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn Gly
785                 790                 795                 800

Xaa Xaa Xaa Xaa Asn Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                805                 810                 815

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            820                 825                 830

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        835                 840                 845

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
    850                 855                 860

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
865                 870                 875                 880

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                885                 890                 895

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            900                 905                 910

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        915                 920                 925
```

<210> 4
<211> 75
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX DNA binding domain (DBD) consensus sequence

<220>
<221> MOD_RES
<222> (1)..(75)
<223> Xaa = any amino acid

<400> 4

```
Thr Leu Gln Trp Leu Xaa Xaa Asn Tyr Xaa Xaa Xaa Glu Gly Val Xaa
  1               5                  10                  15

Leu Pro Arg Xaa Xaa Leu Tyr Xaa His Tyr Leu Xaa Xaa Cys Xaa Xaa
             20                  25                  30
```

```
Xaa Lys Leu Glu Pro Xaa Xaa Ala Ala Xaa Phe Gly Lys Xaa Ile Arg
         35                  40                  45

Xaa Xaa Phe Xaa Xaa Leu Xaa Thr Arg Arg Leu Gly Thr Arg Gly Xaa
     50                  55                  60

Ser Lys Tyr His Tyr Tyr Gly Ile Xaa Xaa Lys
 65                  70                  75


<210> 5
<211> 35
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX B domain
      consensus sequence

<220>
<221> MOD_RES
<222> (1)..(35)
<223> Xaa = any amino acid

<400> 5
Val Xaa Xaa Leu Xaa Xaa Xaa Tyr Xaa Xaa His Cys Xaa Xaa Ile Leu
  1               5                  10                  15

Asp Xaa Xaa Xaa Asn Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
             20                  25                  30

Xaa Phe Trp
         35


<210> 6
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX C domain
      consensus sequence

<220>
<221> MOD_RES
<222> (1)..(40)
<223> Xaa = any amino acid

<400> 6
Leu Tyr Xaa Xaa Xaa Xaa Xaa Xaa Leu Ile Pro Xaa Xaa Xaa Xaa Xaa
  1               5                  10                  15

Xaa Pro Xaa Xaa Leu Xaa Xaa Xaa Ile Arg Xaa Phe Ala Lys Xaa Xaa
             20                  25                  30

Xaa Xaa Trp Xaa Xaa Xaa Xaa Leu
         35                  40
```

<210> 7
<211> 170
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX
dimerization domain consensus sequence

<220>
<221> MOD_RES
<222> (1)..(170)
<223> Xaa = any amino acid

<400> 7
Phe Xaa Xaa Xaa Leu Xaa Arg Xaa Thr Ser Xaa Xaa His Leu Ala Gln
1 5 10 15

Xaa Ala Arg Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Asn Xaa Met Xaa
20 25 30

Ser Asp Xaa Xaa Arg Val Asp Xaa Asn Xaa Xaa Xaa Xaa Gln Ala Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln Xaa Xaa Lys Xaa Xaa Leu Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Glu Xaa Xaa Xaa Glu Trp Leu Asp Xaa Val Xaa Xaa
85 90 95

Gln Xaa Xaa Xaa Xaa Xaa Xaa Tyr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Lys
100 105 110

Xaa Ala Xaa Xaa Phe Leu Leu Lys Trp Ser Phe Xaa Xaa Xaa Xaa Val
115 120 125

Xaa Xaa Xaa Leu Thr Leu Xaa Xaa Ala Xaa Ser Phe Gly Ser Phe His
130 135 140

Leu Ile Arg Xaa Leu Xaa Asp Glu Tyr Xaa Xaa Xaa Xaa Xaa Glu Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu
165 170

<210> 8
<211> 76
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:exemplary RFX
DNA binding domain (DBD)

<400> 8
Thr Leu Gln Trp Leu Glu Glu Asn Tyr Ile Val Cys Glu Gly Val Cys
1 5 10 15

```
Leu Pro Arg Cys Ile Leu Tyr Ala His Tyr Leu Asp Phe Cys Arg Lys
            20                  25                  30

Glu Lys Leu Glu Pro Ala Cys Ala Ala Thr Phe Gly Lys Thr Ile Arg
        35                  40                  45

Gln Lys Phe Pro Leu Leu Thr Thr Arg Arg Leu Gly Thr Arg Gly His
    50                  55                  60

Ser Lys Tyr His Tyr Tyr Gly Ile Gly Ile Lys Glu
65                  70                  75


<210> 9
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:exemplary RFX B
      domain

<400> 9
Lys Val Asp Thr Leu Ile Met Met Tyr Lys Thr His Cys Gln Cys Ile
  1               5                  10                  15

Leu Asp Asn Ala Ile Asn Gly Asn Phe Glu Glu Ile Gln His Phe Leu
            20                  25                  30

Leu His Phe Trp
        35


<210> 10
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:exemplary RFX C
      domain

<400> 10
Leu Tyr Lys Val Leu Thr Asp Val Leu Ile Pro Ala Thr Met Gln Glu
  1               5                  10                  15

Met Pro Glu Ser Leu Leu Ala Asp Ile Arg Asn Phe Ala Lys Asn Trp
            20                  25                  30

Glu Gln Trp Val Val Ser Ser Leu
        35                  40


<210> 11
<211> 179
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:exemplary RFX
      dimerization domain
```

<400> 11
Arg Phe Val Ser Ser Leu Lys Arg Gln Thr Ser Phe Leu His Leu Ala
1 5 10 15

Gln Ile Ala Arg Pro Ala Leu Phe Asp Gln His Val Val Asn Ser Met
20 25 30

Val Ser Asp Ile Glu Arg Val Asp Leu Asn Ser Ile Gly Ser Gln Ala
35 40 45

Leu Leu Thr Ile Ser Gly Ser Thr Asp Thr Glu Ser Gly Ile Tyr Thr
50 55 60

Glu His Asp Ser Ile Thr Val Phe Gln Glu Leu Lys Asp Leu Leu Lys
65 70 75 80

Lys Asn Ala Thr Val Glu Ala Phe Ile Glu Trp Leu Asp Thr Val Val
85 90 95

Glu Gln Arg Val Ile Lys Thr Ser Lys Gln Asn Gly Arg Ser Leu Lys
100 105 110

Lys Arg Ala Gln Asp Phe Leu Leu Lys Trp Ser Phe Phe Gly Ala Arg
115 120 125

Val Met His Asn Leu Thr Leu Asn Asn Ala Ser Ser Phe Gly Ser Phe
130 135 140

His Leu Ile Arg Met Leu Leu Asp Glu Tyr Ile Leu Leu Ala Met Glu
145 150 155 160

Thr Gln Phe Asn Asn Asp Lys Glu Gln Glu Leu Gln Asn Leu Leu Asp
165 170 175

Lys Tyr Met

<210> 12
<211> 76
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX4 DNA
binding domain

<400> 12
Thr Leu Gln Trp Leu Glu Glu Asn Tyr Glu Ile Ala Glu Gly Val Cys
1 5 10 15

Ile Pro Arg Ser Ala Leu Tyr Met His Tyr Leu Asp Phe Cys Glu Lys
20 25 30

Asn Asp Thr Gln Pro Val Asn Ala Ala Ser Phe Gly Lys Ile Ile Arg
35 40 45

Gln Gln Phe Pro Gln Leu Thr Thr Arg Arg Leu Gly Thr Arg Gly Gln
50 55 60

Ser Lys Tyr His Tyr Tyr Gly Ile Ala Val Lys Glu
65 70 75

<210> 13
<211> 77
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX5 DNA
binding domain

<400> 13
Ala Tyr Arg Trp Ile Arg Asn His Leu Glu Glu His Thr Asp Thr Cys
1 5 10 15

Leu Pro Lys Gln Ser Val Tyr Asp Ala Tyr Arg Lys Tyr Cys Glu Ser
20 25 30

Leu Ala Cys Cys Arg Pro Leu Ser Thr Ala Asn Phe Gly Lys Ile Ile
35 40 45

Arg Glu Ile Phe Pro Asp Ile Lys Ala Arg Arg Leu Gly Gly Arg Gly
50 55 60

Gln Ser Lys Tyr Cys Tyr Ser Gly Ile Arg Arg Lys Thr
65 70 75

<210> 14
<211> 75
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:C. elegans RFX
protein daf-19 DNA binding domain

<400> 14
Thr Val Asn Trp Leu Phe Glu Asn Tyr Glu Ile Gly Glu Gly Ser Leu
1 5 10 15

Pro Arg Cys Glu Leu Tyr Asp His Tyr Lys Lys His Cys Ala Glu His
20 25 30

Arg Met Asp Pro Val Asn Ala Ala Ser Phe Gly Lys Leu Ile Arg Ser
35 40 45

Val Phe His Asn Leu Lys Thr Arg Arg Leu Gly Thr Arg Gly Asn Ser
50 55 60

Lys Tyr His Tyr Tyr Gly Ile Arg Leu Lys Asp
65 70 75

<210> 15
<211> 76
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX3 DNA
binding domain

<400> 15
His Leu Gln Trp Leu Leu Asp Asn Tyr Glu Thr Ala Glu Gly Val Ser
1 5 10 15

Leu Pro Arg Ser Thr Leu Tyr Asn His Tyr Leu Arg His Cys Gln Glu
20 25 30

His Lys Leu Asp Pro Val Asn Ala Ala Ser Phe Gly Lys Leu Ile Arg
35 40 45

Ser Ile Phe Met Gly Leu Arg Thr Arg Arg Leu Gly Thr Arg Gly Asn
50 55 60

Ser Lys Tyr His Tyr Tyr Gly Ile Arg Val Lys Pro
65 70 75

<210> 16
<211> 76
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX2 DNA
binding domain

<400> 16
His Leu Gln Trp Leu Leu Asp Asn Tyr Glu Thr Ala Glu Gly Val Ser
1 5 10 15

Leu Pro Arg Ser Ser Leu Tyr Asn His Tyr Leu Arg His Cys Gln Glu
20 25 30

His Lys Leu Asp Pro Val Asn Ala Ala Ser Phe Gly Lys Leu Ile Arg
35 40 45

Ser Val Phe Met Gly Leu Arg Thr Arg Arg Leu Gly Thr Arg Gly Asn
50 55 60

Ser Lys Tyr His Tyr Tyr Gly Ile Arg Leu Lys Pro
65 70 75

<210> 17
<211> 76
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX1 DNA
binding domain

<400> 17
Thr Val Gln Trp Leu Leu Asp Asn Tyr Glu Thr Ala Glu Gly Val Ser
1 5 10 15

Leu Pro Arg Ser Thr Leu Tyr Cys His Tyr Leu Leu His Cys Gln Glu
20 25 30

Gln Lys Leu Glu Pro Val Asn Ala Ala Ser Phe Gly Lys Leu Ile Arg
35 40 45

Ser Val Phe Met Gly Leu Arg Thr Arg Arg Leu Gly Thr Arg Gly Asn
50 55 60

Ser Lys Tyr His Tyr Tyr Gly Leu Arg Ile Lys Ala
65 70 75

<210> 18
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX DNA binding domain consensus sequence

<400> 18
Thr Leu Gln Trp Leu
1 5

<210> 19
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX DNA binding domain consensus sequence

<400> 19
Ala Glu Gly Val Ser Leu Pro Arg Ser
1 5

<210> 20
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX DNA binding domain consensus sequence

<400> 20
Pro Val Asn Ala Ala Ser Phe Gly Lys Leu Ile Arg Ser
1 5 10

<210> 21
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX DNA binding domain consensus sequence

<400> 21
Thr Arg Arg Leu Gly Thr Arg Gly Asn Ser Lys Tyr His Tyr Tyr Gly
1 5 10 15

Ile Arg

<210> 22
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:C. elegans RFX protein daf-19 B domain

<400> 22
Glu Leu Asn Ser Leu Ile Asp Ile Tyr Glu Ile Leu Cys Arg Glu Ile
1 5 10 15

Leu Ala Leu Ile Lys Asn Ile Asp Phe Ala Ser Val Glu Asp Thr Trp
20 25 30

Ser Lys Phe Trp
35

<210> 23
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX1 B domain

<400> 23
Asp Ile Lys Ala Phe Gln Val Leu Tyr Arg Glu His Cys Glu Ala Ile
1 5 10 15

Val Asp Val Met Val Asn Leu Gln Phe Thr Leu Val Glu Thr Leu Trp
20 25 30

Lys Thr Phe Trp
35

<210> 24
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX2 B domain

<400> 24
Asp Val Lys Ala Leu Gln Leu Val Tyr Arg Arg His Cys Glu Ala Thr
1 5 10 15

Val Asp Val Val Met Asn Leu Gln Phe His Tyr Ile Glu Lys Leu Trp
20 25 30

Leu Ser Phe Trp
35

<210> 25
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX4 B domain

<400> 25
Lys Val Ser Thr Phe Ile Met Met Tyr Arg Thr His Cys Gln Arg Ile
1 5 10 15

Leu Asp Thr Val Ile Arg Ala Asn Phe Asp Glu Val Gln Ser Phe Leu
20 25 30

Leu His Phe Trp
35

<210> 26
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX3 B domain

<400> 26
Asp Ile Lys Ser Leu Gln Ser Leu Tyr Arg Glu His Cys Glu Ala Ile
1 5 10 15

Leu Asp Val Val Val Asn Leu Gln Phe Ser Leu Ile Glu Lys Leu Trp
20 25 30

Gln Thr Phe Trp
35

<210> 27
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX4 C domain

<400> 27
Leu Tyr Lys Ala Ile Ser Gly Val Leu Met Pro Thr Val Leu Gln Ala
1 5 10 15

Leu Pro Asp Ser Leu Thr Gln Val Ile Arg Lys Phe Ala Lys Gln Leu
20 25 30

Asp Glu Trp Leu Lys Val Ala Leu
35 40

<210> 28
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX1 C domain

<400> 28
Leu Tyr Gln Gly Leu Val Glu Ile Leu Ile Pro Asp Val Leu Arg Pro
1 5 10 15

Ile Pro Ser Ala Leu Thr Gln Ala Ile Arg Asn Phe Ala Lys Ser Leu
20 25 30

Glu Ser Trp Leu Thr His Ala Met
35 40

<210> 29
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:C. elegans RFX protein daf-19 C domain

<400> 29
Leu Tyr Gln Thr Ile Val Asp Thr Leu Ile Pro Asn Val Leu Leu Ser
1 5 10 15

Glu Leu Ser Thr Gly Met Thr Gln Thr Cys Arg Thr Phe Ala Lys Asn
20 25 30

Ile Asp Val Tyr Leu Arg Lys Ser Leu
35 40

<210> 30
<211> 174
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX4 dimerization domain

<400> 30
Arg Phe Ser Gln Ile Leu Arg Arg Gln Thr Ser Leu Asn His Leu Cys
1 5 10 15

Gln Ala Ser Arg Thr Val Ile His Ser Ala Asp Ile Thr Phe Gln Met
20 25 30

```
Leu Glu Asp Trp Arg Asn Val Asp Leu Asn Ser Ile Thr Lys Gln Thr
         35                  40                  45

Leu Tyr Thr Met Glu Asp Ser Arg Asp Glu His Arg Lys Leu Ile Thr
     50                  55                  60

Gln Leu Tyr Gln Glu Phe Asp His Leu Leu Glu Glu Gln Ser Pro Ile
 65                  70                  75                  80

Glu Ser Tyr Ile Glu Trp Leu Asp Thr Met Val Asp Arg Cys Val Val
                 85                  90                  95

Lys Val Ala Ala Lys Arg Gln Gly Ser Leu Lys Lys Val Ala Gln Gln
            100                 105                 110

Phe Leu Leu Met Trp Ser Cys Phe Gly Thr Arg Val Ile Arg Asp Met
        115                 120                 125

Thr Leu His Ser Ala Pro Ser Phe Gly Ser Phe His Leu Ile His Leu
    130                 135                 140

Met Phe Asp Asp Tyr Val Leu Tyr Leu Leu Glu Ser Leu His Cys Gln
145                 150                 155                 160

Glu Arg Ala Asn Glu Leu Met Arg Ala Met Lys Gly Glu Gly
                165                 170
```

<210> 31
<211> 170
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:C. elegans RFX protein daf-19 dimerization domain

<400> 31

```
Tyr Leu Gln Gln Gly Leu Lys Arg Tyr Thr Ser Leu Asn His Leu Ala
  1               5                  10                  15

His Ala Ser Arg Gly Val Leu Met Lys Pro Glu Gln Val Gln Gln Met
             20                  25                  30

Tyr Gln Asp Tyr Ile Arg Val Asp Ile Asn Thr Val His Gln Gln Ala
         35                  40                  45

Gly Trp Ile Cys Gly Cys Asp Ser Val Met Val His His Val Asn Asn
     50                  55                  60

Ala Phe Lys His Asn Leu Gln Arg Met Ser Ala Met Glu Val Trp Ala
 65                  70                  75                  80

Glu Trp Leu Glu Ser Ile Val Asp Gln Val Leu Ala Lys Tyr His Asp
                 85                  90                  95

Lys Pro Ala Asn Val Ile Ala Asn Val Gly Lys Gln Phe Leu Leu Asn
            100                 105                 110

Trp Ser Phe Tyr Thr Ser Met Ile Ile Arg Asp Leu Thr Leu Arg Ser
        115                 120                 125
```

```
Ala Met Ser Phe Gly Ser Phe Thr Leu Ile Arg Leu Leu Ala Asp Asp
    130                 135                 140

Tyr Met Tyr Tyr Leu Ile Glu Ser Lys Ile Ala Lys Ala Gly Lys Gln
145                 150                 155                 160

Gln Leu Ile Thr Val Ile Arg Ala Asp Lys
                165                 170
```

<210> 32
<211> 168
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX3
dimerization domain

<400> 32

```
Ala Phe Ala Gln Thr Leu Arg Arg Tyr Thr Ser Leu Asn His Leu Ala
  1               5                  10                  15

Gln Ala Ala Arg Ala Val Leu Gln Asn Thr Ser Gln Ile Asn Gln Met
             20                  25                  30

Leu Ser Asp Leu Asn Arg Val Asp Phe Ala Asn Val Gln Glu Gln Ala
         35                  40                  45

Ser Trp Val Cys Gln Cys Asp Asp Asn Met Val Gln Arg Leu Glu Thr
     50                  55                  60

Asp Phe Lys Met Thr Leu Gln Gln Gln Ser Thr Leu Glu Gln Trp Ala
 65                  70                  75                  80

Ala Trp Leu Asp Asn Val Met Met Gln Ala Leu Lys Pro Tyr Glu Gly
                85                  90                  95

Arg Pro Ser Phe Pro Lys Ala Ala Arg Gln Phe Leu Leu Lys Trp Ser
            100                 105                 110

Phe Tyr Ser Ser Met Val Ile Arg Asp Leu Thr Leu Arg Ser Ala Ala
        115                 120                 125

Ser Phe Gly Ser Phe His Leu Ile Arg Leu Leu Tyr Asp Glu Tyr Met
    130                 135                 140

Phe Tyr Leu Val Glu His Arg Val Ala Gln Ala Thr Gly Glu Thr Pro
145                 150                 155                 160

Ile Ala Val Met Gly Glu Val Arg
                165
```

<210> 33
<211> 168
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX1
dimerization domain

<400> 33
Ala Phe Ala Gln Thr Leu Arg Arg Tyr Thr Ser Leu Asn His Leu Ala
1 5 10 15

Gln Ala Ala Arg Ala Val Leu Gln Asn Thr Ala Gln Ile Asn Gln Met
20 25 30

Leu Ser Asp Leu Asn Arg Val Asp Phe Ala Asn Val Gln Glu Gln Ala
35 40 45

Ser Trp Val Cys Arg Cys Glu Asp Arg Val Val Gln Arg Leu Glu Gln
50 55 60

Asp Phe Lys Val Thr Leu Gln Gln Gln Asn Ser Leu Glu Gln Trp Ala
65 70 75 80

Ala Trp Leu Asp Gly Val Val Ser Gln Val Leu Lys Pro Tyr Gln Gly
85 90 95

Ser Ala Gly Phe Pro Lys Ala Ala Lys Leu Phe Leu Leu Lys Trp Ser
100 105 110

Phe Tyr Ser Ser Met Val Ile Arg Asp Leu Thr Leu Arg Ser Ala Ala
115 120 125

Ser Phe Gly Ser Phe His Leu Ile Arg Leu Leu Tyr Asp Glu Tyr Met
130 135 140

Tyr Tyr Leu Ile Glu His Arg Val Ala Gln Ala Lys Gly Glu Thr Pro
145 150 155 160

Ile Ala Val Met Gly Glu Phe Ala
165

<210> 34
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX
dimerization domain consensus sequence

<400> 34
Leu Arg Arg Tyr Thr Ser Leu Asn His Leu Ala Gln Ala Ala Arg
1 5 10 15

<210> 35
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX
dimerization domain consensus sequence

<400> 35
Asn Gln Met Leu Ser Asp
1 5

<210> 36
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX
dimerization domain consensus sequence

<400> 36
Trp Ala Glu Trp Leu Asp
1 5

<210> 37
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX
dimerization domain consensus sequence

<400> 37
Gln Phe Leu Leu Lys Trp Ser Phe Tyr
1 5

<210> 38
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX
dimerization domain consensus sequence

<400> 38
Ser Met Val Ile Arg Asp Leu Thr Leu Arg Ser Ala
1 5 10

<210> 39
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX
dimerization domain consensus sequence

<400> 39
Ser Phe Gly Ser Phe His Leu Ile Arg Leu Leu
1 5 10

<210> 40
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:RFX
dimerization domain consensus sequence

<400> 40
Asp Glu Tyr Met
1

<210> 41
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:hexahistidine
(His) affinity tag

<400> 41
His His His His His His
1 5

<210> 42
<211> 200
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:poly-Gly
flexible linker

<220>
<221> MOD_RES
<222> (6)..(200)
<223> Gly residues from position 6 to 200 may be present
or absent

<400> 42
Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
1 5 10 15

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
20 25 30

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
35 40 45

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
50 55 60

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
65 70 75 80

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
85 90 95

```
Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
            100                 105                 110

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
        115                 120                 125

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
    130                 135                 140

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
145                 150                 155                 160

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
                165                 170                 175

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
            180                 185                 190

Gly Gly Gly Gly Gly Gly Gly Gly
        195                 200
```